



# Surprisal, the PDC, and the primary locus of processing difficulty in relative clauses

Roger Levy<sup>1\*</sup> and Edward Gibson<sup>2,3</sup>

<sup>1</sup> Department of Linguistics, University of California, San Diego, San Diego, CA, USA

<sup>2</sup> Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

<sup>3</sup> Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA, USA

\*Correspondence: [rlevy@ucsd.edu](mailto:rlevy@ucsd.edu)

Edited by:

Charles Clifton, University of Massachusetts Amherst, USA

Of the ambitious purview of MacDonald's (2013) article, we find the part fleshed out in most concrete detail—the comprehension consequences of her Production-Distribution-Comprehension (PDC) theory, the easiest to comment upon. Such a theory as she has sketched out would be extraordinarily compelling: a theory that, in contrast with accounts relying on “innate parsing biases,” posits that “comprehension results reflect distributional regularities in the language” that “comprehenders are generating expectations for upcoming input,” places “emphasis on the role of learning probabilistic constraints,” makes use of “extensive language corpora” to “[permit] comprehension researchers to examine the relationship between production patterns ... and comprehension behavior” and thereby “reframes our understanding of sentence comprehension.” The only way we can see such a theory being more compelling would be for it to be specified precisely enough to be computationally implementable and to make *quantitative* and *localized* predictions about the processing difficulty of every word in a sentence that could be tested rigorously on a variety of linguistic materials. A naïve reader of MacDonald's article may not know that such a theory already exists and has been highly successful. This theory, known as *surprisal*, was first proposed by Hale (2001), building on early ideas by Attneave (1959) from the dawn of information theory (Shannon, 1948) and cognitive science.

As proposed by Hale (2001) and elaborated by Levy (2008), surprisal theory posits that comprehenders use fine-grained probabilistic knowledge derived from linguistic experience to form expectations both about the structural

interpretation of what has already been encountered in the input and about what input may yet be upcoming, and that these expectations immediately determine processing difficulty (with a precise, quantitative difficulty metric) and guide interpretation preferences. The theory has been applied to a variety of languages and linguistic phenomena, it has been tested on comprehension behavior of both specific grammatical constructions (e.g., Brouwer et al., 2010; Levy et al., 2012) and naturalistic datasets (Boston et al., 2008; Demberg and Keller, 2008), and the functional form of its incremental difficulty metric has been empirically confirmed (Smith and Levy, 2008, 2013).

In the first empirical case discussed by MacDonald, surprisal theory predicts the local interpretation preference for precisely the reasons articulated by MacDonald for her PDC theory. Moreover, the Distribution-Comprehension (DC) part of MacDonald's theory—the idea that the empirical distribution of various syntactic and semantic properties of language determines probabilistic (hence defeasible) processing preferences—is explicit in models predating surprisal, including not only the constraint-based approaches she mentions but also in the probabilistic parsing approach of Jurafsky (1996); and since distribution can only be derived from production, it seems to us that the Production-Distribution (PD) part is implicit. The theoretical advance of surprisal over these earlier probabilistic and constraint-based approaches is very specific: it unified probabilistic resolution of structural ambiguity already present in the input with the formation of expectations regarding future input. It is unclear

what corresponding conceptual advance is provided by MacDonald's account.

This brings us to the second empirical case of comprehension behavior discussed by MacDonald: the processing of relative clauses (RCs). We deeply appreciate the point that the relative production frequencies of subject and object RCs are highly sensitive to a variety of factors including (but not limited to) language, NP type (e.g., pronominal vs. full; Real and Christiansen, 2007), and agent and patient animacy (e.g., Gennari et al., 2012). We also agree that an impressive body of research points to the generalization that the comparative processing difficulty of subject vs. object RCs is often well-predicted by their relative frequencies given these factors (Traxler et al., 2002; Real and Christiansen, 2007; Gennari and MacDonald, 2008; Levy et al., 2013; cf. Fedorenko et al., 2011; Gibson et al., in press; and see also Doyle and Levy, 2010; Gibson and Wu, 2013 for counter-examples). Nevertheless, we do not believe that any extant theory in the class MacDonald proposes adequately explains all the critical facts in the syntactic complexity of relative clauses. In particular, the critical data bear not only on which types of RCs are hardest to comprehend, but also on the *locus* of maximal processing difficulty. This point is extremely clear for the classic SRC/ORC processing difficulty differential for English:

- (1a) The reporter that attacked the senator admitted the error. (SRC)
- (1b) The reporter that the senator attacked admitted the error. (ORC)

A surfeit of theories—both experience- and memory-based—correctly predict

that the ORC *that the senator attacked* in (1b) is harder than the SRC *that attacked the senator* in (1a). However, fully incremental experience-based theories such as surprisal fail to correctly predict *where* difficulty arises in (1b). As pointed out by Hale (2001), Grodner and Gibson (2005), and Levy (2008), experience-based theories predict the locus of processing difficulty for ORCs to be at the onset of the RC noun phrase *the senator*, which in (1b) disconfirms the comprehender's rational expectation that the RC will turn out to be subject-extracted. Although recent work (Staub, 2010) has revealed some degree of processing difficulty at this point in ORCs like (1b), the bulk of the difficulty clearly falls on the RC verb *attacked*. This difficulty occurs despite the fact that the RC verb in ORCs should be more expected, not less expected, than in SRCs: there are fewer syntactic events that can occur after *The reporter that the senator ...* than after *The reporter that ...*, and there are fewer things that a senator can do to a reporter than that a reporter can do (see discussion in Grodner and Gibson, 2005 and Levy, 2008). These are precisely the conditions under which surprisal correctly predicts difficulty differentials for a variety of languages and constructions (Konieczny, 2000; Vasishth and Lewis, 2006; Nakatani and Gibson, 2008; Levy and Keller, 2013), yet surprisal fails in the case of English RCs (and it seems to fail similarly for Russian RCs; Levy et al., 2013). Crucially, this problem for surprisal and similar experience-based theories arises regardless of whether one considers the preceding context, *The reporter that ...*, to be structurally "ambiguous": as MacDonald suggests, it is clear that there is considerable indeterminacy as to how the sentence will continue at this point, including indeterminacy as to the grammatical role of the head noun. The key point is that where most of this indeterminacy is pruned away—at the RC noun phrase onset—is not where the differential difficulty is largest<sup>1</sup>.

<sup>1</sup> Hale (2003, 2006) presents an alternative experience-based theory of incremental processing difficulty, the Entropy Reduction Hypothesis, that purportedly predicts the SRC/ORC processing difficulty differential with success. However, we do not believe that this proposal is ultimately empirically viable (see e.g., discussion in Levy et al., 2013).

We consider theories of syntactic processing making reference to explicit, costly (and/or potentially fallible) memory operations, such as those of Gibson (1998, 2000) and Lewis and colleagues (2005, 2006), of continued importance in the study of RC comprehension because they make the right predictions not only about *what* is difficult but about *where* the difficulty is observed in this heavily studied empirical domain.

Although MacDonald's proposal in its present form has not made theoretical commitments as precise as those of surprisal, it is not clear how her proposal could be cashed out to make precise predictions about *where* processing difficulty occurs in a way that avoids the same empirical difficulties that surprisal runs into. This is not to say that there is no hope for developing purely experience-based theories of processing difficulty that explain currently problematic data such as those we describe above. But we do not believe that any such theory currently exists, and we are not sure how to develop one ourselves.

## ACKNOWLEDGMENTS

We would like to thank Ev Fedorenko for her comments.

## REFERENCES

- Attneave, F. (1959). *Applications of Information Theory to Psychology: A Summary of Basic Concepts, Methods and Results*. New York, NY: Holt, Rinehart and Winston.
- Boston, M. E., Hale, J. T., Kliegl, R., Patil, U., and Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: an evaluation using the Potsdam sentence corpus. *J. Eye Mov. Res.* 2, 1–12.
- Brouwer, H., Fitz, H., and Hoeks, J. C. (2010). "Modeling the noun phrase versus sentence coordination ambiguity in Dutch: evidence from surprisal theory," in *Proceedings of the 2010 Workshop on Cognitive Modeling and Computational Linguistics* (Uppsala: Association for Computational Linguistics), 72–80.
- Demberg, V., and Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition* 109, 193–210.
- Doyle, G., and Levy, R. (2010). "A puzzle regarding relative pronoun choice: when frequency and difficulty disagree," in *Poster Presentation Given at the 2010 CUNY Sentence Processing Conference* (New York, NY), March 18–20.
- Fedorenko, E., Tily, H., and Gibson, E. (2011). "A comprehensive investigation of animacy effects in relative clauses," in *The 24th CUNY Conference on Human Sentence Processing* (Palo Alto CA), (Accessed March 2011).

- Gennari, S. P., and MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *J. Mem. Lang.* 58, 161–187.
- Gennari, S. P., Mirković, J., and MacDonald, M. C. (2012). Animacy and competition in relative clause production: a cross-linguistic investigation. *Cogn. Psychol.* 65, 141–176.
- Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition* 68, 1–76.
- Gibson, E. (2000). "The dependency locality theory: a distance-based theory of linguistic complexity," in *Image, Language, Brain*, eds A. Marantz, Y. Miyashita, and W. O'Neil (Cambridge, MA: MIT Press), 95–126.
- Gibson, E., Tily, H., and Fedorenko, E. (in press). "The processing complexity of English relative clauses," in *Language Down the Garden Path: The Cognitive and Biological Basis for Linguistic Structure*, eds M. Sanz, I. Laka, and M. Tanenhaus. (Oxford: Oxford University Press).
- Gibson, E., and Wu, I. (2013). Processing Chinese relative clauses in context. *Lang. Cogn. Process.* 28, 125–155.
- Grodner, D., and Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cogn. Sci.* 29, 261–291.
- Hale, J. (2001). "A probabilistic Earley parser as a psycholinguistic model," in *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics* (Pittsburgh, PA), 159–166.
- Hale, J. (2003). The information conveyed by words in sentences. *J. Psychol. Res.* 32, 101–123.
- Hale, J. (2006). Uncertainty about the rest of the sentence. *Cogn. Sci.* 30, 609–642.
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cogn. Sci.* 20, 137–194.
- Konieczny, L. (2000). Locality and parsing complexity. *J. Psychol. Res.* 29, 627–645.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition* 106, 1126–1177.
- Levy, R., Fedorenko, E., and Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *J. Mem. Lang.* (in press).
- Levy, R., Fedorenko, E., Breen, M., and Gibson, E. (2012). The processing of extraposed structures in English. *Cognition* 122, 12–36.
- Levy, R., and Keller, F. (2013). Expectation and locality effects in German verb-final structures. *J. Mem. Lang.* 68, 199–202.
- Lewis, R. L., and Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cogn. Sci.* 29, 1–45.
- Lewis, R. L., Vasishth, S., and Van Dyke, J. (2006). Computational principles of working memory in sentence comprehension. *Trends Cogn. Sci.* 10, 447–454.
- MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Front. Psychol.* 4:226. doi: 10.3389/fpsyg.2013.00226
- Nakatani, K., and Gibson, E. (2008). Distinguishing theories of syntactic expectation cost in sentence comprehension: evidence from Japanese. *Linguistics* 46, 63–87.
- Real, F., and Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *J. Mem. Lang.* 57, 1–23.

- Shannon, C. (1948). A mathematical theory of communications. *Bell Syst. Tech. J.* 27, 623–656.
- Smith, N. J., and Levy, R. (2008). “Optimal processing times in reading: a formal model and empirical investigation,” in *Proceedings of the 30th Annual Meeting of the Cognitive Science Society* (Washington, DC).
- Smith, N. J., and Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition* (in press).
- Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition* 116, 71–86.
- Traxler, M. J., Morris, R. K., and Seely, R. E. (2002). Processing subject and object relative clauses: evidence from eye movements. *J. Mem. Lang.* 47, 69–90.
- Vasishth, S., and Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and anti-locality effects. *Language* 82, 767–794.
- Received: 26 February 2013; accepted: 11 April 2013; published online: 14 May 2013.
- Citation: Levy R and Gibson E (2013) Surprisal, the PDC, and the primary locus of processing difficulty in relative clauses. *Front. Psychol.* 4:229. doi: 10.3389/fpsyg.2013.00229
- This article was submitted to *Frontiers in Language Sciences*, a specialty of *Frontiers in Psychology*. Copyright © 2013 Levy and Gibson. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.